Identification And Expression Analysis of *FORMIN* Gene Family In Rice (*Oryza Sativa*)

Bin Li

Chengdu University of Technology

Zhi-Ye Du

Chengdu University of Technology

Shan He

Chengdu University of Technology

Kai Xiao

Chengdu University of Technology

Xing Wang

Chengdu University of Technology

Kai-Shun Wang

Chengdu University of Technology

Ji Chen

Sichuan Agricultural University

Jin Huang (✉ huangjin18@cdut.edu.cn)

Chengdu University of Technology

Research Article

Keywords: FORMINs, bioinformatic, rice, gene structure, expression profiles.

Posted Date: December 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1132176/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: FORMIN proteins, which are composed of proteins containing FH1 and FH2 domains, play crucial roles in the growth and development of organisms. However, the functions of FORMINs in rice (*Oryza sativa*) remain largely unclear.

Results: In this study, a total of 17 *FORMIN* genes were identified in rice, *OsFH17* was the first time identified in this study. In addition, the distribution on chromosomes, gene structure, as well as conserved motifs of rice *FORMIN*s was investigated. According to their protein structural and phylogenetic features, these 17 rice *FORMIN* genes were classified into two distinct subfamilies. Subcellular localization prediction showed that rice FORMINs were located in cytosol, golgi, endoplasmic reticulum, extracellular, and vacuole. Protein protein interaction (PPI) prediction results shown that FORMIN protein might answer hormone signals and be involved in cytoskeleton dynamics regulation and cell wall morphology regulation. The results of silico analysis and qRT-PCR confirmation of the gene expression showed that the expression of rice *FORMIN*s were related to their tissue distribution. Moreover, *OsFH3*, *OsFH5* and *OsFH7* were upregulated under phytohormone treatments.

Conclusions: Overall, our research may shed light on the understanding and further investigation of the biological functions of rice *FORMIN*s.

Background

Actins play critical roles in cellular activities such as cytoskeleton construction, cell division, and cell polarized growth[1–7]. As actin-binding proteins (ABPs), FORMINs function as actin-filament nucleators that bind to the barbed ends of actin filaments to promote the polymerization and elongation of actin filaments[8]. Therefore, FORMINs are extremely paramount to regulate cell polarized growth, cytokinesis, and cell migrations [1, 5, 9, 10].

Plant FORMINs are much different from the ones in animals or fungi[11, 12]. For example, FORMINs in plants do not have GTPase binding domain (GBD), diaphanous inhibitory domain (DID), or diaphanous auto-regulatory domains (DAD)[13]. Instead, phosphatase tensin (PTEN) domain could be found in some plant FORMIN proteins but not in animal or fungal FORMIN proteins[10, 14]. Besides, comparing to the GTPase binding and activation of the animal FORMIN model, how plant FORMINs are activated remains unclear, as no GTPase binding domain has been found in plant FORMINs yet[15].

Although some studies have been done to identify the members or to explore the functions of *FORMIN* genes, The knowledge of plant *FORMIN* genes is far to be clear[16]. For example, it has been reported that there are 21 and 16 *FORMIN* genes that have been identified in arabidopsis and rice genomes respectively[13]. But only 3 of the 16 rice *FORMIN* genes have been experimentally explored for their functions. *OsFH1* is involved in the polarized elongation of root hairs, loss-of-function of *OsFH5* gene resulted in irregular cell growth in leaf sheath and shorter cells in inner husks, and *OsFH15* is involved in seed size regulation by interacting with actin filaments and microtubules[17–19]. Therefore, the functions of rice *FORMIN* genes are far to be clear. To further elucidate the functions of rice FORMINs, both bioinformatics and molecular biological strategies were employed in this study. The cis-acting elements, temporal and spatial expression pattern and the conserved protein domains of rice *FORMIN* genes or proteins were analyzed along with the prediction analysis of the proteins that interact with rice FORMIN proteins. Our study may provide new insights for understanding how FORMINs function during rice growth and development.

Results

Identification and Phylogenetic Analysis of FORMIN Genes

We used the Formin homology 2 (FH2) profile of FH2 domain (PF02181) from Pfam (<https://pfam.xfam.org/>) database as reference sequences to extract rice FORMIN family members by using TBtools and GenomeNet with default parameters. The FORMIN proteins of *Cyanidioschyzon merolae* (red alga), *Saccharomyces cerevisiae* (baker's yeast), *Physcomitrella patens* (*P. patens*), *Selaginella moellendorffii* (spike moss), *Medicago truncatula* (barrel medic), *Lotus japonicus* (*L. japonicus*), *Zea mays* (maize) and *Arabidopsis thaliana* (arabidopsis) were downloaded from genome database and verified by TBtools software with default parameters. A total of 17 proteins were identified belonging to the *FORMIN* gene family in rice, *OsFH17* was the first time identified in this study. The results shown that the number of rice FORMIN's amino acids ranged from 364 to 1511 (Table 1). Furthermore, we identified 3, 2, 8, 5, 15, 14, 20 and 21 FORMIN proteins in red alga, baker's yeast, *P. patens*, spike moss, barrel medic, *L. japonicus*, maize and arabidopsis, respectively. In order to explore the evolutionary relationships of the *FORMIN* gene family, an unrooted phylogenetic tree was constructed using the full length FORMIN protein sequences from these nine selected species (Figure 1). Obviously, all 105 proteins were divided into two major clades, clade I including type I FORMIN subfamily and clade II including type II FORMIN subfamily (Figure 1). Regarding the 17 rice FORMIN proteins, they were divided into two types (type I, type II) as well. Type I FORMIN family consists of *OsFH3*, *OsFH5*, *OsFH6*, *OsFH7* and *OsFH12*. Type II consists of *OsFH1*, *OsFH2*, *OsFH4*, *OsFH8*, *OsFH9*, *OsFH10*, *OsFH11*, *OsFH13*, *OsFH14*, *OsFH15*, *OsFH16* and *OsFH17* (Figure 1).

Interestingly, type II FORMIN proteins were not found in red alga, baker's yeast and spike moss. Besides, *P. patens* contained 6 type II FORMIN proteins. Moreover, rice, barrel medic, *L. japonicus*, maize and arabidopsis all contained over 6 type II FORMIN proteins. In addition, the survey of eumycophyta and plant genomes for *FORMIN* genes indicated that the numbers of type I and type II *FORMIN* genes in the angiosperm were more than the ones in gymnosperm and eumycophyta (Figure 2). These results imply that *FORMIN* genes had been diverged during evolution, probably deriving from family expansion and domain acquisition.

Gene structure or conserved domains of FORMIN Genes or FORMIN proteins

The analyses of gene structure and gene phylogenetic relation could help us to better understand the function or evolution of less explored genes. Therefore, an NJ tree was generated with MEGA7 using all the FORMIN protein sequences from rice genome, and the structure of these genes was constructed as well (Figure 3). As shown in Figure 3, the distribution of exons and introns of rice was variable, and the number of exons and the length of sequences of these

FORMIN genes were widely ranged between two subclasses. Type I *FORMIN* subfamily includes 2 genes with 1 exon, 8 genes with 3 exons, 1 gene with 4 exons and 1 gene with 5 exons. Whereas the type II includes 1 gene with 5 exons, 1 gene with 12 exons, 1 gene with 13 exons and 2 genes with 15 exons (Table 1).

Besides the study of gene structure features, the full-length sequences of rice *FORMIN* proteins were submitted to MEME website for architecture analysis. The results indicated that ten conserved motifs were identified in these rice *FORMIN* family proteins, and the differences of motifs between type-I and type-II *FORMIN*s were significant. Such as motif 5 and motif 9 were unique to type-I, *FORMIN*s while motif 10 was specific to all type-II *FORMIN*s. Meanwhile, in the same subfamily, the same type and distribution mode of motifs were shared, indicating that the protein architecture of these rice *FORMIN* proteins were highly conserved within a specific subfamily.

Table 1
The gene and protein features of *FORMIN* family members in rice

Gene name	RAP-ID	Chromosome localization	CDS length (bp)	Exon number	Protein length (aa)	Protein MW (kDa)	Isoelectric Point (pI)	Subcellular localization	GOterms
<i>OsFH1</i>	<i>Os01g0897700</i>	Chr1	2883	3	961	101642	9.1195	cytosol	C:extracellular
<i>OsFH2</i>	<i>Os04g0461800</i>	Chr4	2502	3	834	88899.1	8.5935	golgi	C:extracellular
<i>OsFH3</i>	<i>Os10g0119300</i>	Chr10	1161	13	387	43312	5.9619	cytosol	C:nucleus
<i>OsFH4</i>	<i>Os10g0347800</i>	Chr10	2550	3	850	90280.9	9.7161	cytosol	C:extracellular
<i>OsFH5</i>	<i>Os07g0596300</i>	Chr7	1914	4	638	70922.8	4.3667	cytosol	C:chloroplast thylakoid lumen
<i>OsFH6</i>	<i>Os08g0280200</i>	Chr8	2682	15	894	95782.6	9.2704	cytosol	C:nucleus
<i>OsFH7</i>	<i>Os02g0794900</i>	Chr2	1092	12	364	41156.7	7.7994	cytosol	C:nucleus
<i>OsFH8</i>	<i>Os03g0204100</i>	Chr3	1629	1	543	55704.4	10.7337	extracellular	C:extracellular
<i>OsFH9</i>	<i>Os08g0431200</i>	Chr8	2649	3	883	97860.1	8.6902	cytosol	C:extracellular
<i>OsFH10</i>	<i>Os02g0161100</i>	Chr2	2646	5	882	96463.3	5.7992	golgi	C:extracellular
<i>OsFH11</i>	<i>Os07g0545500</i>	Chr7	2790	5	930	100990	8.0425	golgi	C:extracellular
<i>OsFH12</i>	<i>Os04g0245000</i>	Chr4	4533	15	1511	165148	5.4272	cytosol	C:nucleus
<i>OsFH13</i>	<i>Os07g0588200</i>	Chr7	2325	3	775	83810.3	8.7229	vacuole	C:extracellular
<i>OsFH14</i>	<i>Os05g0104000</i>	Chr5	2493	3	831	90221.6	8.4292	endoplasmic reticulum	C:nucleus
<i>OsFH15</i>	<i>Os09g0517600</i>	Chr9	2367	3	789	84580.3	8.8049	endoplasmic reticulum	C:extracellular
<i>OsFH16</i>	<i>Os02g0739100</i>	Chr2	2721	1	907	95565.3	10.3198	endoplasmic reticulum	C:extracellular
<i>OsFH17</i>	<i>Os04g0100300</i>	Chr4	2721	3	907	95753.4	8.338	golgi	C:extracellular

In addition, the conserved FH2 domains in all rice *FORMIN* proteins were also analyzed by using multiple sequence alignment, and the results showed that all conserved motifs were present in FH2 domains (Figure 3c). It is noteworthy that *OsFH12* (belonging to type II *FORMIN* subfamily) contained a PTEN structure in its N-terminus (Figure 3C). It was also found that these 17 rice *FORMIN*s were not evenly distributed on the 12 chromosomes of rice (Figure 3D). Twelve of the rice *FORMIN* genes were localized on the first ten chromosomes, including Chr 1, Chr 2, Chr 3, Chr 4, Chr 5, Chr 7, Chr 8, Chr 9 and Chr 10. None of the rice *FORMIN*s was found on the chromosome 6, 11 or 12 and no *FORMIN* gene cluster was found on rice chromosomes (Figure 3D).

Protein-protein Interaction Prediction

To determine the functions of rice *FORMIN* proteins, protein–protein interaction (PPI) prediction was performed. Among 17 rice *FORMIN* proteins, three of them interacted with cell wall morphology regulation-related proteins and proteins involved in auxin (IAA) signal pathways. Totally 11 rice *FORMIN* proteins were closely connected with actin polymerization-related proteins, but only *OsFH5* interacted with actin-depolymerizing-related proteins. Except for *OsFH2* and *OsFH3*, other rice *FORMIN*s showed a close relationship with cell cytoskeleton-regulated proteins, which implies the importance of *FORMIN*s in regulating cytoskeleton dynamics in plants (Table 2). Some protein-protein interaction prediction results were also consistent with previous studies that rice *FORMIN*s are involved in cell elongation, cell wall loosening and assembly of microtubules and microfilaments[17, 19, 20]. Moreover, the interaction prediction between *FORMIN* family members also indicated that there might be interactions between *OsFH10*, *OsFH2*, *OsFH14*, *OsFH11*, *OsFH8* and *OsFH15* (see Supplementary Figure S3, Supplementary Table S2). Such as, *OsFH11* was predicted to interact with *OsFH14*, *OsFH2*, *OsFH15*, or *OsFH8*. *OsFH10*, *OsFH14* and *OsFH2* were predicted to interact with each other.

Table 2
Analysis of protein-protein interaction of rice FORMINs

Name	Formins	Actin related	Microtubule Related	Cytoskeleton Structure Related	Cell Wall Formation Related	Exocytic Vesicles Plasma Membrane Docking Related	Kinesin-like Protein	Transcription factor	Silicon Efflux Transporter	K ⁺ Channel Related	Auxin Signal Pathway
<i>OsFH1</i>	OsFH5	-	KIN5A	profilin LP04	CESA1, EXPA17, BC1L4	-	-	-	-	-	IAA3
<i>OsFH2</i>	OsFH16	Act	-	-	-	SEC10	-	-	-	-	-
<i>OsFH3</i>	OsFH14	-	-	-	-	-	-	SH3, TFIIA γ	-	-	-
<i>OsFH4</i>	OsFH16	ARP2	-	profilin LP04, profilin A	-	SEC10	-	-	-	-	-
<i>OsFH5</i>	OsFH1	ARP2, ADF9	-	profilin LP04, profilin A	-	-	-	-	-	-	BG1
<i>OsFH6</i>	-	ARP2	-	profilin LP04, profilin A	-	-	-	-	-	-	-
<i>OsFH7</i>	-	ARP2	-	profilin LP04, profilin A	-	-	-	-	-	-	-
<i>OsFH8</i>	OsFH16	ARP2	-	profilin LP04, profilin A	-	SEC10	-	-	-	-	-
<i>OsFH9</i>	-	ARP2	TUBG2	profilin LP04, profilin A	-	-	KIN13B	-	-	-	-
<i>OsFH10</i>	OsFH16	ARP2	-	profilin LP04, profilin A	-	-	-	-	-	-	-
<i>OsFH11</i>	OsFH12	ARP2	TUBG2	profilin LP04, profilin A	-	-	-	-	-	-	-
<i>OsFH12</i>	OsFH11	Act	-	profilin LP04, profilin A	-	-	-	-	LSI3	-	-
<i>OsFH13</i>	OsFH16	ARP2	-	profilin LP04, profilin A	-	SEC10	-	-	-	-	-
<i>OsFH14</i>	OsFH3	-	KIN5A	profilin LP04, profilin A	-	-	-	-	-	FIP2	-
<i>OsFH15</i>	-	ARP2	-	profilin LP04, profilin A	-	-	-	-	-	FIP2	BG1
<i>OsFH16</i>	OsFH2, OsFH4, OsFH8, OsFH13	-	-	profilin LP04, profilin A	-	SEC10	-	-	-	-	-
<i>OsFH17</i>	OsFH16	ARP2	-	profilin LP04, profilin A	-	-	-	-	-	-	-

Expression profiles of rice FORMIN genes in different tissues or in response to phytohormone treatments

To understand the temporal and spatial expression patterns of different *FORMIN* genes, microarray data from the rice XPro database was used to assess the expression profile of these *FORMIN* genes across different tissues or organs of rice plants (anther, callus, leaf, panicle, pistil, root, shoot and seed), or under a number of phytohormone treatments (abscisic acid, gibberellin, IAA, brassinosteroid, cytokinin and jasmonic acid). Meanwhile, the tissue specificity of the expression of *FORMIN* genes was also validated through quantitative real time-PCR (qRT-PCR) by using the cDNA from different tissues or organs (whole seedling, shoot, shoot apices, leaf blade, leaf sheath, root, immature flower and mature flower).

The expression analysis data have shown that all these 17 genes were significantly differentially expressed in selected tissues or organs (anther, callus, leaf, panicle, pistil, root, shoot and seed). *OsFH1*, *OsFH16*, *OsFH5*, *OsFH13*, *OsFH5* and *OsFH3* showed relatively higher but variable expression levels in almost all

selected tissues or organs. The following four genes had relatively higher expression in some specific tissues or organs: (i) *OsFH17* in the anthers, (ii) *OsFH14* and *OsFH15* in the calli, and (iii) *OsFH7* in the roots.

The qRT-PCR analysis results showed that all these 17 genes could be divided into five groups, which include cluster 1 with three genes (*OsFH5*, *OsFH1* and *OsFH2*), cluster 2 with two genes (*OsFH7* and *OsFH10*), cluster 3 with one gene (*OsFH6*), cluster 4 with six genes (*OsFH9*, *OsFH12*, *OsFH3*, *OsFH17*, *OsFH5*, *OsFH16*, *OsFH15*) and cluster 5 with five genes (*OsFH4*, *OsFH11*, *OsFH14*, *OsFH8* and *OsFH13*) (Figure 4B). From cluster 1 to cluster 4, they exhibited specific expression in the tissues or organs such as immature flowers, shoot apices and leaves. And the genes of cluster 5 showed expression in leaves, immature flowers and mature flowers.

Phytohormones play crucial roles during rice growth and development. Coincidentally, actin as a downstream element of auxin to regulate plant morphological changes has been elucidated[21]. To verify whether rice *FORMIN*s were regulated by IAA or other phytohormones, the expression profiles of rice *FORMIN*s were systematically established by using microarray data from the RiceXPro. As shown in Figure 4C and D, most rice *FORMIN* genes were ubiquitously but differentially expressed under phytohormone treatments. For instance, *OsFH3* was induced by all phytohormones in the shoots. *OsFH5* was significantly induced by jasmonic acid (MeJA), gibberellin (GA), IAA and abscisic acid (ABA) in the shoots. *OsFH9* showed higher expression levels when treated with cytokinin (CTK) but was repressed by brassinosteroid (BR), GA, IAA and MeJA. However, it has been noticed that some rice *FORMIN* genes showed different expression patterns in roots and shoots. Such as, the expression of *OsFH6* was induced by BR in the roots, but it was repressed by GA, IAA, ABA and MeJA in the shoots. *OsFH7* showed positive expression patterns when treated with ABA, GA and MeJA, but it was repressed by IAA in the roots. In the roots, *OsFH10* was upregulated under GA treatment but showed negative expression patterns when treated with IAA, MeJA or BR.

Discussion

In this study, a total of 17 *FORMIN* family genes from rice have been identified. Syntenic analysis of the *FORMIN* gene family in species revealed that *FORMIN* genes in the plant kingdom were increased during evolution, showing that these genes played a critical role in plants during evolution. As most of the Ka/Ks ratios were less than 1.0, suggesting the purifying selection within the gene family. All these 17 *FORMIN* genes were divided into two subfamilies based on their protein structure. Meanwhile, structural analysis of the *FORMIN* gene family showed that just 1 gene (*OsFH12* of type II) contained PTEN. As research has shown, PTEN is a potent tumour suppressor found in human cancers[22]. Moreover, the PTENs form dimers in order to bind to the cell membrane[23]. For the plants, the studies of PTEN in moss have shown that the PTEN domain is able to bind to PI(3,5)P2 to recruit For2 to the plasma membrane, and located in discrete cortical dots[24]. *OsFH12* contain a PTEN domain, which implies that *OsFH12* may possess some function which is mediated by the ability of lipid-binding.

The molecular weights and the isoelectric point (pI) of these 17 *FORMIN* proteins were different (Table 1), which implied that the difference of rice *FORMIN*s in the physical and chemical properties. Based on the MultiLoc analysis, we found that the *FORMIN* family genes were localized in various intracellular compartments, including endosomes, cytosol, golgi bodies, endoplasmic reticulum, extracellular space and the vacuoles (Table 1). As can be seen, the localization of *OsFH5* is in the range of previously reported experimental results, while the results of *OsFH1* did not agree with the previous reports[25]. It was shown earlier that *OsFH1* was located on the plasma membrane, but in this research was reported located in the cytosol[17]. We could not exclude the possibility that these PM-located *FORMIN*s are recruited to different sites by other proteins or that *FORMIN*s may exhibit variable subcellular localization while at different development stages or under different environmental stimuli. But these results remind us that we cannot ignore the unreliability of the predicted results by bioinformatics method, which may be caused by the complexity of what happens in the cells.

Gene duplication is an important mechanism for increasing genetic variability and creating novel genes. In plants, this mechanism is a significant force in increasing gene family size and the functional expansion of genomes. In our study, a total of 14, 6, 6, or 2 homolog gene pairs were identified between rice, maize, barrel medic, *L. japonicus* and arabidopsis, respectively (Supplementary Figure S2). These results indicated that there was a strong genetic relationship between *FORMIN* genes in angiosperm. However, according to colinear analysis, there was no collinearity between rice with other species.

Studies have shown that the rice morphometric variation depends on the dynamic changes of the cytoskeleton[10, 26]. Actually, genetic association studies also identified *FORMIN*'s crucial role in cytoskeleton regulation, such as *Bni1p* in baker's yeast and *AtFH5* in arabidopsis[27, 28]. Moreover, gene expression patterns can provide important insights into the function of plant *FORMIN* genes. In our study, rice *FORMIN* gene family members showed differential expression patterns, no matter based on the data from the rice XPro database or qRT-PCR analysis. XPro database showed that except *OsFH16*, *OsFH17*, *OsFH3* and *OsFH15* which were highly expressed in differential tissues, other rice *FORMIN* genes showed low expression levels in all tissues. qRT-PCR results showed that most genes had higher expression in leaf (*OsFH3*, *OsFH4*, *OsFH8*, *OsFH9*, *OsFH11*, *OsFH12*, *OsFH13*, *OsFH14*, *OsFH15*, *OsFH16* and *OsFH17*). The expression levels of *OsFH3* and *OsFH13* in the qRT-PCR analysis were consistent with *in silico* analysis results, with both high expressions in leaf. Recently, the AtROP6 have been demonstrated to control interdigitated growth of pavement cells to form jigsaw-puzzle shapes[29, 30]. The ROP-*FORMIN* regulation pathways have been clearly clarified in baker's yeast[27]. Thus those genes which were highly expressed in the leaf may be involved in the pavement cells development. Concurrently, *OsFH6* was not expressed in leaf but in leaf sheath, which also suggested the potential function of *OsFH6*. The function of *FORMIN*s in the pollen tube has been elucidated in arabidopsis, such as *AtFH3* and *AtFH5* are involved in the regulation of apical actin polymerization and actin array construction in pollen tubes[31]. Our qRT-PCR results showed that *OsFH1* and *OsFH2* were highly expressed in the immature flower. However, comparing to the mature flower, *OsFH1* and *OsFH2* were relatively less expressed in other tissues or organs, implying that *OsFH1* and *OsFH2* might play crucial roles in the process of rice flower development.

Nowadays, studies show that actin dynamics is also regulated by phytohormones such as auxin[21]. Thus, plant *FORMIN*s may be involved in this process as well. Our results showed that the expression patterns of *OsFH3*, *OsFH5* and *OsFH9* were altered under phyto-hormones treatments in the shoots. The expression patterns of *OsFH6*, *OsFH7* and *OsFH10* changed under phyto-hormones treatment in the root. Among the *FORMIN* genes altered in shoot, the

function of OsFH5 has been demonstrated in early studies[25]. Studies show that *OsFH5* participates in the regulation of rice morphogenesis by modulating the expansion of diffusely growing cells[20]. In contrast, reported studies have also suggested that IAA regulates rice cell expansion[32]. Moreover, the cis-element analysis showed that the upstream sequence of *OsFH5* included IAA-related cis-elements (Supplementary Figure S1). Consequently, *OsFH5* was likely involve in the regulation of IAA-mediated cell expansion. Further, through gene promoter element analysis, the upstream sequence of *OsFH3* included ABA, MeJA and GA related cis-elements. The *OsFH7* also contained MeJA and GA related cis-elements, suggesting that the expression of *OsFH3* and *OsFH7* might be regulated by phytohormones. Additionally, some environmental stress-related cis-acting elements, like drought-responsive elements, stress-responsive elements and low-temperature-responsive elements, were also found in the rice *FORMIN* promoter regions. These results indicated that rice *FORMIN* genes might be closely related to the responses to multiple abiotic stresses. Moreover, all rice *FORMIN*s contained a number of light-responsive elements, suggesting that rice *FORMIN* genes are important players in the light response of rice.

Overall, in this study, we systematically identified rice *FORMIN* genes in the rice genome, established a foundation for further investigation of rice *FORMIN* genes, and provided novel insights into their biological functions.

Methods

Identification of FORMIN Sequences

Rice sequence-related information (gene sequences, reference genome, protein sequences, annotation and features) was retrieved from the Ensembl plants database (<https://plants.ensembl.org/index.html>)[33]. To identify FORMIN proteins in rice, we downloaded the FORMIN homology 2 domain proteins sequences from Pfam databases as reference sequences. Then the Formin proteins were extracted from the rice genome database by using TBtools (version 1.089) and GenomeNet with default parameters[34]. For phylogenetic analysis, the FORMIN proteins sequences of baker's yeast was downloaded from the saccharomyces genome database (SGD) (<https://www.yeastgenome.org/>)[35]. The red alga, *P. patens*, spike moss, barrel medic, *L. japonicus*, maize and arabidopsis were downloaded from the GreenPhyl database (<https://www.greenphyl.org/>)[36]. All FORMIN proteins used for phylogenetic analysis were verified through alignment against the FORMIN homology 2 domain proteins sequences.

Sequence Analysis

The length of sequences, molecular weights (MW), and isoelectric point (pI) of FORMIN genes were extracted from RAP-DB (<https://rapdb.dna.affrc.go.jp/>)[37]. The rice *FORMIN* genes IDs was used to predict the GOterms of Fomrin genes through the online bioinformatics tool GENE ONTOLOGY (<http://geneontology.org/>)[38]. Furthermore, the subcellular localization of rice *FORMIN* genes were predicted by MultiLoc (<http://www-bs.informatik.uni-tuebingen.de/Services/MultiLoc/>)[39]. Then the MEME website (<https://meme-suite.org/>) was used to identify the conserved motifs of *FORMIN* genes[40]. To obtain the cis-acting elements information in rice *FORMIN*s promoters, 2-kb upstream sequences of each rice *FORMIN* genes were intercepted and submitted to the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>) for further analysis[41].

Multiple Alignments And Phylogenetic Analysis

For phylogenetic analysis, rice FH2 domain sequences of candidate proteins were extracted and aligned with the FH2 domain sequences of red alga, baker's yeast, *P. patens*, spike moss, barrel medic, *L. japonicus*, maize and arabidopsis using ClustaW. The evolutionary tree was constructed by MEGA7.0 with Neighbor-Joining (NJ) method (1000 bootstrap replications) and visualized by using the iTOL (<http://itol.embl.de/>)[42].

Chromosome Localization and Gene Duplication Analysis

The physical location information of *rice FORMIN* genes in the chromosomes was obtained from the genome annotation files of rice. The genome sequences information of baker's yeast, *P. patens*, barrel medic, *L. japonicus*, maize and arabidopsis *FORMIN* was downloaded from the Ensembl plants database. Besides, the duplication events within the subgenome was detected by the Multiple Collinearity Scan toolkit (MCScanX) with default settings[43]. Then the TBtools were used to exhibit the chromosomes distribution and the synteny relationship of orthologous *FORMIN* genes that identified from baker's yeast, *P. patens*, barrel medic, *L. japonicus*, maize and arabidopsis. Furthermore, KaKs Calculator 2.0 was employed to estimate the non-synonymous (Ka) and synonymous (Ks) substitution rates of the duplicated *FORMIN* genes.

Expression Analysis of rice FORMIN Genes in silico

The public available genes expression data of rice in different tissues and in response to different phytohormones were downloaded from the RED database (<http://expression.ic4r.org/>) and the rice Xpro database (<https://ricexpro.dna.affrc.go.jp/>)[44, 45]. Then the values were performed treatment of log2 to construct the heatmap by the TBtools.

Expression Analysis of rice FORMIN Genes by qRT-PCR

Oryza sativa spp. Japonica cultivar 'Dongjin' was used in this study. The full shape and disease-free seeds were selected and planted on one layer of vinyl mesh that touched the surface of the 1/2 MS solutions (pH 5.8). The growth conditions were set at 28°C with 16 h light/8 h dark cycle.

After 14 days of cultivation, the plant's whole seedling, shoot, root, shoot apices were selected as samples. After 42 days of cultivation, leaf blade and leaf sheath were selected as samples. Moreover, 80 days cultivated immature flower and 100 days cultivated mature flower also selected as samples.

Total RNA of the samples was extracted using the Plant RNA miniprep kit. Then the RevertAid First Strand cDNA Synthesis Kit (Thermo) was used to synthesize the first strand of cDNA from the high-quality DNafree RNA. qRT-PCR was performed by using qTOWER3G instrument with SYBR Green. The qRT-PCR reaction conditions were as follows: 95°C for 2 min; 40 cycles of 95°C/15 s, 55°C/15 s, 72°C/20 s. Every treatment contained three biological replications, and each analysis was repeated by three times. The data from qRT-PCR was analyzed using the $2^{-\Delta\Delta Ct}$ method, and the expression levels were represented by the mean values of the three replicates. The student's t test was carried out to determine whether the changes in gene expression were significant. The heatmap of correlations was generated by TBtools. And sequences of the primers used in this study were shown in detail in Supplementary Table S1.

Availability of data and materials

The FORMIN homology 2 domain (PF02181) sequences were downloaded from the Pfam (<https://pfam.xfam.org/>). The FORMIN amino acid sequences of alga, *P. patens*, spike moss, barrel medic, *L. japonicus*, maize and arabidopsis were downloaded from GreenPhyl (<https://www.greenphyl.org/>). The FORMIN amino acid sequence of baker's yeast was downloaded from the *saccharomyces* genome database (SGD) (<https://www.yeastgenome.org/>). Rice sequence-related information (gene sequences, reference genome sequences, protein sequences, annotation and features) was retrieved from Ensembl plants database (<https://plants.ensembl.org/index.html>). The information of sequences length, molecular weight (MW), and isoelectric point (pI) of *FORMIN* genes and proteins were extracted from RAP-DB (<https://rapdb.dna.affrc.go.jp/>). The evolutionary tree was visualized by using iTOL (<http://itol.embl.de/>). The GO terms of *FORMIN* genes were analyzed by the online bioinformatics tool GENE ONTOLOGY (<http://geneontology.org/>). The subcellular localization of rice *FORMIN* genes were predicted by MultiLoc (<http://www.bs.informatik.uni-tuebingen.de/Services/MultiLoc/>). The conserved motifs of *FORMIN* genes were predicted by MEME website (<https://meme-suite.org/>). To obtain the cis-acting elements information of rice *FORMINs* promoters, 2-kb upstream sequences of each rice *FORMIN* genes were intercepted and submitted to PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>). The publicly available expression data of rice genes were downloaded from RED database (<http://expression.ic4r.org/>) and the rice Xpro database (<https://ricexpro.dna.affrc.go.jp/>). Public access to all databases is open. All data generated or analyzed during this study are included in this published article (and its Additional files).

Abbreviations

ABPs: actin-binding proteins; GBD: GTPase Binding Domain; DID: Diaphanous Inhibitory Domain; DAD: Diaphanous Auto-regulatory Domain; PTEN: Phosphatase Tensin; Red alga (Cm): *Cyanidioschyzon merolae*; *P. patens* (Pp): *Physcomitrella patens*; Spike moss (Sm): *Selaginella moellendorffii*; Barrel medic (Mt): *Medicago truncatula*; *L. japonicus* (Lj): *Lotus japonicus*; Maize (Zm): *Zea mays*; Arabidopsis (At): *Arabidopsis thaliana*; Rice (Os): *Oryza sativa*; IAA: auxin; MeJA: jasmonic acid; GA: gibberellin; ABA: abscisic acid; CTK: cytokinin; BR: brassinosteroid;

Declarations

Funding

This work was supported by the National Natural Science Foundation of China (Grant Nos. 31870383).

Author Contributions

HJ, and CJ conceived and designed the experiments and provided funding for research work. LB, DZY and HS performed most of the experiments. XK, WX and WKS performed some of the experiments. BL and SH wrote the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgements

At the point of finishing this paper, we sincerely thank all those have lent me hands while my writing this paper.

Ethics approval and consent to participate

Not applicable.

References

1. Campellone KG, Welch MD: A nucleator arms race: cellular control of actin assembly. *Nat Rev Mol Cell Biol* 2010, 11(4):237–251.
2. Pollard TD, Cooper JA: Actin, a central player in cell shape and movement. *Science* 2009, 326(5957):1208–1212.
3. Mathur J, Hülskamp M: Microtubules and microfilaments in cell morphogenesis in higher plants. *Curr Biol* 2002, 12(19):R669-676.
4. Chen J, Wang F, Zheng S, Xu T, Yang Z: Pavement cells: a model system for non-transcriptional auxin signalling and crosstalks. *J Exp Bot* 2015, 66(16):4957–4970.
5. Zhang Y, Dong J: Cell polarity: compassing cell division and differentiation in plants. *Curr Opin Plant Biol* 2018, 45(Pt A):127–135.
6. Wu J, Akhmanova A: Microtubule-Organizing Centers. *Annu Rev Cell Dev Biol* 2017, 33:51–75.
7. Lian G, Kanaujia S, Wong T, Sheen V: FilaminA and Formin2 regulate skeletal, muscular, and intestinal formation through mesenchymal progenitor proliferation. *PLoS One* 2017, 12(12):e0189285.
8. Martin SG, Chang F: Dynamics of the Formin For3p in Actin Cable Assembly. *Current Biology* 2006, 16(12):1161–1170.
9. Young KG, Copeland JW: Formins in cell signaling. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research* 2010, 1803(2):183–190.
10. Vidali L, Gisbergen P, Guérin C, Franco P, Li M, Burkart G, Augustine R, Blanchoin L, Bezanilla M: Rapid formin-mediated actin-filament elongation is essential for polarized plant cell growth. *Proceedings of the National Academy of Sciences of the United States of America* 2009, 106:13341–13346.

11. Deeks MJ, Hussey PJ, Davies B: Formins: intermediates in signal-transduction cascades that affect cytoskeletal reorganization. *Trends Plant Sci* 2002, 7(11):492–498.
12. Liu R, Linardopoulou EV, Osborn GE, Parkhurst SM: Formins in development: orchestrating body plan origami. *Biochim Biophys Acta* 2010, 1803(2):207–225.
13. Cvrčková F, Novotný M, Pícková D, Zárský V: Formin homology 2 domains occur in multiple contexts in angiosperms. *BMC Genomics* 2004, 5(1):44–44.
14. Kollárová E, Baquero Forero A, Stillerová L, Přerostová S, Cvrčková F: Arabidopsis Class II Formins AtFH13 and AtFH14 Can Form Heterodimers but Exhibit Distinct Patterns of Cellular Localization. *International Journal of Molecular Sciences* 2020, 21(1).
15. Liu W, Sato A, Khadka D, Bharti R, Diaz H, Runnels LW, Habas R: Mechanism of activation of the Formin protein Daam1. *Proceedings of the National Academy of Sciences* 2007.
16. Higgs HN, Peterson KJ: Phylogenetic analysis of the formin homology 2 domain. *Mol Biol Cell* 2005, 16(1):1–13.
17. Huang J, Kim CM, Xuan Y, Liu J, Kim T, Kim B-K, Han C-D: Formin homology 1 (OsFH1) regulates root-hair elongation in rice (*Oryza sativa*). *Planta* 2013, 237.
18. Yi-Ran TAO Y-ZX, Jia XIE,Wei-Jiang TIAN,Xiao-Qiong ZHANG,Xiao-Bo ZHANG,Qian ZHOU,Xian-Chun SANG,Xiao-Wen WANG: Identification and Gene Mapping of *sdb1* Mutant with a Semi-dwarfism and Bigger Seed in Rice. *Acta Agronomica Sinica* 2018, 44(11):1621–1630.
19. Sun T, Li S, Ren H: OsFH15, a class I formin, interacts with microfilaments and microtubules to regulate grain size via affecting cell expansion in rice. *Scientific Reports* 2017, 7(1):6538.
20. Wang J, Xue X, Ren H: New insights into the role of plant formins: regulating the organization of the actin and microtubule cytoskeleton. *Protoplasma* 2012, 249(2):101–107.
21. Scheuring D, Löfke C, Krüger F, Kittelmann M, Eisa A, Hughes L, Smith RS, Hawes C, Schumacher K, Kleine-Vehn J: Actin-dependent vacuolar occupancy of the cell determines auxin-induced growth repression. *Proceedings of the National Academy of Sciences* 2016, 113(2):452-457.
22. Lee Y-R, Chen M, Pandolfi PP: The functions and regulation of the PTEN tumour suppressor: new modes and prospects. *Nature Reviews Molecular Cell Biology* 2018, 19(9):547–562.
23. Das S, Dixon JE, Cho W: Membrane-binding and activation mechanism of PTEN. *Proceedings of the National Academy of Sciences of the United States of America* 2003, 100(13):7491–7496.
24. Gisbergen P, Li M, Wu S-Z, Bezanilla M: Class II formin targeting to the cell cortex by binding PI(3,5)P2 is essential for polarized growth. *J Cell Biol* 2012, 198:235–250.
25. Chang S, Yang L, Liang W, Xie Q, Huang G, Wang D, Zhu W, Zhang D: RFPs Interact with OsFH5 and Control the Rice Morphology by Affecting Stability of OsFH5; 2021.
26. Frank MJ, Smith LG: A small, novel protein highly conserved in plants and animals promotes the polarized growth and division of maize leaf epidermal cells. *Curr Biol* 2002, 12(10):849–853.
27. Liu W, Santiago-Tirado FH, Bretscher A: Yeast formin Bni1p has multiple localization regions that function in polarized growth and spindle orientation. *Molecular biology of the cell* 2012, 23(3):412–422.
28. Ingouff M, Fitz Gerald J, Guérin C, Robert H, Sørensen M, Van Damme D, Geelen D, Blanchoin L, Berger F: Plant formin AtFH5 is an evolutionarily conserved actin nucleator involved in cytokinesis. *Nature cell biology* 2005, 7:374–380.
29. Fu Y, Xu T, Zhu L, Wen M, Yang Z: A ROP GTPase Signaling Pathway Controls Cortical Microtubule Ordering and Cell Expansion in Arabidopsis. *Current Biology* 2009, 19(21):1827–1832.
30. Lin D, Ren H, Fu Y: ROP GTPase-mediated auxin signaling regulates pavement cell interdigitation in Arabidopsis thaliana: ROP Signaling Regulates Pavement Cell Interdigitation. *Journal of Integrative Plant Biology* 2014, 57.
31. Lan Y, Liu X, Fu Y, Huang S, Yang Z: Arabidopsis class I formins control membrane-originated actin polymerization at pollen tube tips. *Plos Genetics* 2018, 14(11).
32. Li G, Liang W, Zhang X, Ren H, Hu J, Bennett MJ, Zhang D: Rice actin-binding protein RMD is a key link in the auxin–actin regulatory loop that controls cell growth. *Proceedings of the National Academy of Sciences* 2014, 111(28):10377.
33. Howe KL, Contreras-Moreira B, De Silva N, Maslen G, Akanni W, Allen J, Alvarez-Jarreta J, Barba M, Bolser DM, Cambell L *et al*: Ensembl Genomes 2020—enabling non-vertebrate genomic research. *Nucleic Acids Research* 2019, 48(D1):D689-D695.
34. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, Xia R: TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol Plant* 2020, 13(8):1194–1202.
35. Skrzypek MS, Nash RS: Biocuration at the Saccharomyces genome database. *Genesis* 2015, 53(8):450–457.
36. Valentin G, Abdel T, Gaëtán D, Jean-François D, Matthieu C, Mathieu R: GreenPhyloDB v5: a comparative pangenomic database for plant genomes. *Nucleic Acids Res* 2021, 49(D1):D1464-d1471.
37. Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, Wakimoto H, Yang CC, Iwamoto M, Abe T *et al*: Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol* 2013, 54(2):e6.
38. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT *et al*: Gene Ontology: tool for the unification of biology. *Nature Genetics* 2000, 25(1):25–29.
39. Höglund A, Dönnies P, Blum T, Adolph H-W, Kohlbacher O: MultiLoc: prediction of protein subcellular localization using N-terminal targeting sequences, sequence motifs and amino acid composition. *Bioinformatics* 2006, 22(10):1158–1165.

40. Bailey TL, Johnson J, Grant CE, Noble WS: The MEME Suite. *Nucleic Acids Research* 2015, 43(W1):W39-W49.

41. Lescot M, Déhais P, Thijs G, Marchal K, Moreau Y, Van de Peer Y, Rouzé P, Rombauts S: PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res* 2002, 30(1):325–327.

42. Letunic I, Bork P: Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Research* 2021, 49(W1):W293-W296.

43. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, Lee TH, Jin H, Marler B, Guo H *et al*: MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res* 2012, 40(7):e49.

44. Xia L, Zou D, Sang J, Xu X, Yin H, Li M, Wu S, Hu S, Hao L, Zhang Z: Rice Expression Database (RED): An integrated RNA-Seq-derived gene expression database for rice. *J Genet Genomics* 2017, 44(5):235–241.

45. Sato Y, Takehisa H, Kamatsuki K, Minami H, Namiki N, Ikawa H, Ohyanagi H, Sugimoto K, Antonio BA, Nagamura Y: RiceXPro version 3.0: expanding the informatics resource for rice transcriptome. *Nucleic Acids Res* 2013, 41(Database issue):D1206-1213.

Figures

Figure 1

Bio Neighbor-Joining (BioNJ) phylogenetic tree (1000 bootstraps), based on FH2 domain of FORMIN genes from baker's yeast, red alga, P. patens, spike moss, barrel medic, L. japonicus, maize, arabidopsis and rice. The abbreviations of the species are as follows: Cm Cyanidioschyzon merolae (red alga), Pp Physcomitrella patens (P. patens), Sm Selaginella moellendorffii (spike moss), Mt Medicago truncatula (barrel medic), Lj Lotus japonicus (L. japonicus), Zm Zea may (maize), At Arabidopsis thaliana (arabidopsis) and Os Oryza sativa (rice).

Figure 2

A survey of FORMIN proteins across the plant kingdom and eumycophyta. performed by using ensemblplants databases. For confirmed FORMIN proteins, the FH2 domains were further screened using AtFHs sequence as query primary sequence and TBtools as the search tool with default parameters. Numbers in parentheses indicate total FORMIN proteins produced by the indicated plant species. The taxonomic hierarchy of the selected plant species surveyed was shown at up.

Figure 3

The gene structure and motif composition of rice FORMIN genes, and chromosomal mapping of FORMIN genes in rice. (3A) OsFHs phylogenetic tree of rice. The tree was prepared based on the FH2 domains using the Clustal X algorithm. The confidence of groupings was estimated using 1000 bootstrap replicates. (3B) The conserved motif analysis of rice FORMIN proteins. (3C) The conserved domain of OsFHs. (3D) Chromosomal mapping of OsFHs in rice.

Figure 4

Expression profiles of rice FORMIN genes in different tissues and their response to phytohormone treatment. (4A) Expression profiles of rice FORMIN genes in 8 different tissues (anther, callus, leaf, panicle, pistil, root, shoot and seed). The data used in the figure were obtained from the Xpro database. (4B) Gene expression profiling of the rice FORMIN genes in eight different tissues by qRT-PCR (whole seedling, shoot, shoot apices, leaf blade, leaf sheath, root, immature flower and mature flower). (4C) Expression profiles of OsFHs in rice shoot under different phytohormone treatments. The data used in the figure were obtained from the Xpro database. (4D) Expression profiles of OsFHs in rice root under different phytohormone treatments. The data used in the figure were obtained from the Xpro database. All red represents up-regulated genes, and blue is for down-regulated ones.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [AdditionalFigures.pptx](#)
- [AdditionalTable1.xlsx](#)
- [AdditionalTable2.xlsx](#)